DOCUMENT RESUME

ED 051 455 AC 010 348

AUTHOR	McGuigan, F. J.
TITLE	How To Select and Evaluate Programmed Instructional Materials.
INSTITUTION	North Carolina State Univ., Raleigh. School of Education.
SPONS AGENCY	Bureau of Adult, Vocational , and Technical Education (DHEW/OE), Washington, D.C.
REPORT NO	Occas-Pap-1
BUREAU NO	BR-4-32028
PUB DATE	Apr 71
GRANT	OEG-0-70-4599(324)
NOTE	20p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement Tests, Data Analysis, *Evaluation Techniques, *Instructional Materials, Learning Activities, *Programed Instruction, *Program Evaluation

ABSTRACT

The large number of programs commercially available at all educational levels presents the teacher with a major selection problem. The purpose of this discussion is to explain the assessment procedures that publishers should use in order to assure the publication of only high quality programs and to share this information with educators. When he first receives a program, a publisher should have it checked over by an expert, and a decision should be rendered as to whether or not it is academically sound and whether it is appropriate for a specific educational curriculum. The program is then objectively tested in a classroom situation. Once appropriate sample subjects are selected, the students are administered an objective achievement test as a pre-test. Next, they work through their programs. Attempts are made to minimize the effects on the students' learning of all variables other than the program itself. Upon completion of the program, the students take the achievement test again as a post-test. Finally, the students and teachers complete standard program evaluation forms. The data gathered from the pre- and post-tests, from the students' actual work on the program, and from the evaluation forms are then analyzed. A final publish-no publish decision is made according to the criterion of the amount a student learns from the use of the program. (CK)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

HOW TO SELECT AND EVALUATE PROGRAMMED INSTRUCTIONAL MATERIALS

ED051455

F. J. McGUIGAN

APRIL, 1971

ALC

ADULT LEARNING CENTER

School of Education, North Carolina State University, Raleigh, North Carolina

AC 010 348

1

ADULT LEARNING CENTER

The Adult Learning Center was approved and established as an experimental and demonstration project in adult basic education in 1967 under the provisions of Section 309(b) of the Adult Education Act of 1966. The Center is an integral part of the research and development program of the School of Education at North Carolina State University. The purpose of the project is the identification, development, and evaluation of innovative curriculum materials and instructional methods that will accelerate and enhance learning in undereducated adults. Special emphasis has been placed upon investigations into the utilization of modern educational technology for the instruction of undereducated adults. A major concern of the Center is the development of packaged instructional materials and improved instructional methods which are capable of being institutionalized within adult basic education programs in public school systems. The Center is totally committed to the belief that its research and developmental efforts must give promise of materially improving the ongoing adult basic education program at all levels, local, state, and national.

The Center is part of the program conducted under the Auspices of the Bureau of Adult, Vocational, and Technical Education, Division of Adult Education Programs, Office of Education, U. S. Department of Health, Education, and Welfare. The program of the Center cuts across the Schools of Education, Agriculture and Life Sciences, Liberal Arts, and Physical Sciences and Applied Mathematics at North Carolina State University at Raleigh.

THE OCCASIONAL PAPER SERIES

The Occasional Paper Series has been established by the Center as a vehicle for the publication of monographs of general interest to people involved in all areas of adult education. For the most part, these monographs consist of papers or speeches written or delivered by members of the Center staff, as a consequence of their affiliation with the Center.

ADDITIONAL INFORMATION

For additional information regarding the Center, please write to:

Dr. D. Barry Lumsden, Director
Adult Learning Center
P. O. Box 5096
North Carolina State University at Raleigh
Raleigh, North Carolina 27607

ED051455

HOW TO SELECT AND EVALUATE PROGRAMMED INSTRUCTIONAL MATERIALS

F. J. McGUIGAN, SPECIAL CONSULTANT
ADULT LEARNING CENTER
NORTH CAROLINA STATE UNIVERSITY AT RALEIGH

* *

This paper is published pursuant to a grant from the Office of Education, U. S. Department of Health, Education, and Welfare. Grantees undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

* *

Occasional Paper No. 1

U. S. Department of Health, Education, and Welfare
Office of Education
Bureau of Adult, Vocational, and Technical Education
Division of Adult Education Programs
Project No. 432028, Grant No. OEG-0-70-4599 (324)

1971

PREFACE

In recent years there has occurred a rather dramatic proliferation of commercially-produced programmed instructional materials. It is frequently the case that programmed instructional materials are represented as the panacea to the achievement problems of students of all age and ability levels. Because programmed materials allow students to progress at their own rate of speed, underachievers are guaranteed eventual success in the content mastery of their programs. For the same reason, brighter students are allowed to progress as rapidly as their abilities allow inasmuch as group instructional restraints have been removed from the learning situation.

It is indeed unfortunate that in all too many cases there is little or no evidence that particular programmed instructional materials are effective; it is not known whether or not they teach and, if so, under what conditions. The effectiveness of a programmed instructional package is directly proportional to the extent to which it teaches; learning outcomes provide the best measures of the effectiveness of programmed materials.

In this report, Dr. McGuigan sets forth several guidelines which should be of assistance to those educators whose concern and responsibility is the identification, selection, and testing of programmed texts. This report is not presented as the ultimate and only statement on the subject. It is sent forth, however, with the confidence that it constitutes an important contribution to the accumulating

body of knowledge on programmed instruction. Its obvious utility is the Center's justification for its inclusion in the Occasional Paper Series.

In addition to serving as a special consultant to the Adult Learning Center, Dr. McGuigan is Professor of Psychology at Hollins College, Roanoke, Virginia.

D. Barry Lumsden, Director
Adult Learning Center

HOW TO SELECT AND EVALUATE PROGRAMMED INSTRUCTIONAL MATERIALS

The large number of programs commercially available at all educational levels presents the teacher with a major selection problem--which should she choose for her students? By sharing with the teacher the validating data that led to his decision to publish a given program, the publisher can help the teacher answer her question. Some publishers make this type of information available in separate evaluation reports of the programs that it markets. The purpose of the present discussion is to explain the assessment procedures that publishers should use in order to assure publication of only high quality programs and to share this information with educators to the benefit of their students. We shall now consider the various criteria for assessing the quality of any given program.

Pre-classroom Testing Criteria

Programming Technique. Programmed learning has resulted in many products of widely varying programming quality. When a publisher first receives a program, it should be checked over by a programming expert to ascertain whether or not sound principles of programming technique have been used by the author. Only those programs that have superior programming characteristics should be further considered.

Subject-Matter-Expert Evaluation. The next step is to have the program thoroughly studied by one or more subject-matter experts. A decision is rendered as to whether or not the program is academically sound, whether or not the terms, principles, mechanics, etc. are

up-to-date, and so forth. Finally, the subject-matter expert decides whether or not the content of the program is appropriate for a specific educational curriculum. Providing that the program passes these first two tests, it is then objectively tested in a classroom situation.

Assessment In A Classroom

Efforts are made to select samples that are representative of the population for which the program is intended. Hence, appropriate grade levels and classes are chosen, the I.Q.'s of the "experimental" students are ascertained to make sure that they are nationally typical, and so forth. See Appendix A for an example of a form which might be used for the collection of important student and student performance information.

Once appropriate samples are selected, the students are administered an objective achievement test as a pre-test. Following this they work through their programs, usually at their own rate. During this learning period attempts are made to minimize the effects on the students' learning of all variables other than the program itself, e.g., the activities of the teacher are restricted to procedural matters and the students work on their programs in the classroom under the teacher's supervision. Once the students have completed their programs, they take the achievement test again as a post-test. Finally, the students and the teachers complete standard program evaluation forms (See Appendices B and C).

The data gathered from the pre- and post-tests, from the student's actual work in his program, and from the evaluation forms are then analyzed. A final publish-no publish decision is made according to the

following criteria. These criteria are guidelines that the teacher might profitably use in helping her decide whether or not to adopt a given program.

Classroom Testing Criteria

Learning Data. The primary purpose of the program is to teach--if the student does not learn a substantial amount from the use of the program, the other criteria may be disregarded. The question, then, of how much students learn from the use of a program is much more important than all of the other criteria combined.

To assess amount learned as accurately as possible we have developed a ratio which we call G (McGuigan and Peters, 1965). To better understand G let us refer to Table 1. Suppose that it is possible to score 100 points

Table 1

Learning Data Analysis (Means)

Possible Score	=	100
Pre-Test Score	=	20
Post-Test Score	=	85
Gain Score	=	85 - 20 = 65
Possible Gain Score	=	100 - 20 = 80
$G = \frac{\text{Gain Score}}{\text{Possible Gain Score}}$	=	$\frac{65}{80} = .81$

on the achievement test; in this case, the possible score is, obviously 100. Now suppose that a class makes an average score of 20 before the

program was taken (the pre-test score) and of 85 upon completion of the program (the post-test score). The gain score can thus be seen to be 85 minus 20, or 65. Next we ascertain the possible gain score, i.e., the maximum amount of gain that the students could have demonstrated. This value is the possible score (100) minus the pre-test score (20), which is 80. Now, to compute the ratio, G, one merely needs to divide the possible gain score (80) into the actual gain score (65), resulting in this example in the value .81.

With this understanding of how to compute the gain-to-possible-gain ratio, G, let us now consider how this value helps us to answer our major question of whether or not the students learned a sufficient amount from the use of any given program. Put otherwise: is any given value of G (e.g., .81) sufficiently high that we can conclude that the program led to an adequate amount of learning? To answer this question, we have constructed a frequency distribution of G scores for a number of programs that have been tested to date (Figure 1). Note that G typically varies between 0.0 and 1.0 such that the higher the G value, the greater the learning. By studying the G scores that make up the frequency distribution presented in Figure 1 we can see, for instance, that one program yielded a value within the .10 to .19 category, that two programs had G values of .20 to .29, and so forth. Now, to answer our question, note that the G value for our hypothetical program is located within this frequency distribution by means of the unshaded region. Since this value of .81 is a high one, relative to G resulting from the other programs tested, we may conclude that this program led to a relatively superior amount of learning.

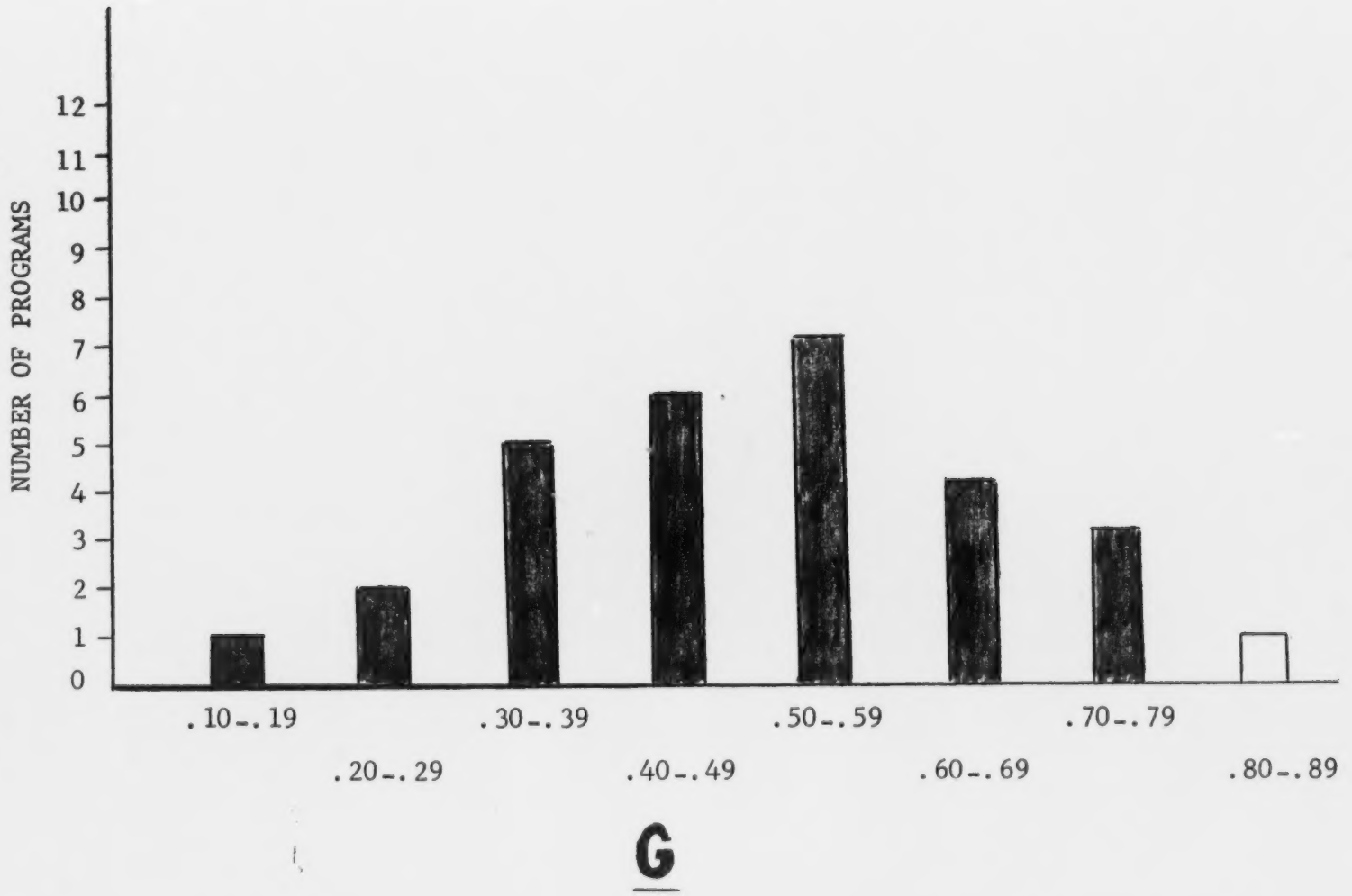

FIG. 1. Frequency distribution of gain to possible gain ratios ($\underline{G}$) for 28 programs previously tested. The non-shaded region locates the $\underline{G}$ value for the test program at a very high position relative to the other programs tested. (The 28 programs previously tested were not tested by the author under the aegis of the Adult Learning Center.)

Student Evaluations. It has been shown that student reactions are positively related to amount learned (McGuigan and Peters, 1965) and, in any event, one should prefer using programs that students regard as enjoyable and beneficial. The standard student evaluation form includes six items on which students rate the program. In using this criterion, we compare the student reactions to a test program with all student reactions to previous programs that have been tested. That is, we sum the student reactions for all programs tested to date and then compare their reactions to the program currently undergoing testing (McGuigan and Peters, 1965). Consider Table 2 as an example. We can see that 68% of the 1489 students who studied a variety of programs indicated that they learned more because a program was used in their course. In contrast, 81% said that they learned more because the test program was used. We may thus regard the student reactions to the test program in this example as being relatively favorable. Each report of the validating data for any given program presents these kinds of findings in detail.

Table 2

Student Reactions To Test Program

Question #1: Because a program was used in this course, I believe:

	N	I learned more	It made no difference	I learned less
31 Programs	1489	68%	19%	13%
Test Program	59	81%	7%	12%

Teacher Evaluation. The teachers' reactions, as recorded on standard evaluation forms, are verbally summarized in each report.

Error Analysis. Research has also indicated that the higher the error rate, the less the learning (McGuigan and Peters, 1965). This finding is consistent with principles of learning (especially of the Skinnerian variety). Hence, both empirical and theoretical considerations indicate that the error rate for a given program should be relatively low. To ascertain the mean error rate, the responses that the students wrote in their programs are carefully checked and the number of errors is counted for each student. The mean number of errors is then determined for the entire sample of students, and this value is divided by the total number of responses called for by the program. The resulting value is the mean error rate. Figure 2 presents a frequency distribution of mean error rates for programs that have been previously tested. Figure 2 can now be used to determine whether or not the error rate for a given program is excessive. For example, suppose that a program yielded a mean error rate of 2.3%. This value, represented by the unshaded region in Figure 2, can be seen to be a low one, relative to the mean error rates for the entire sample of programs.

Concluding Statement

In summary, then, one who is considering adopting any given program should first request and examine the learning data made available by the publisher. Program and learning data obtained from publishers should, as much as possible, conform to the Recommendations for Reporting the Effectiveness of Programed Instruction Material prepared by the Joint Committee on Programmed Instruction and Teaching Machines. If the information obtained

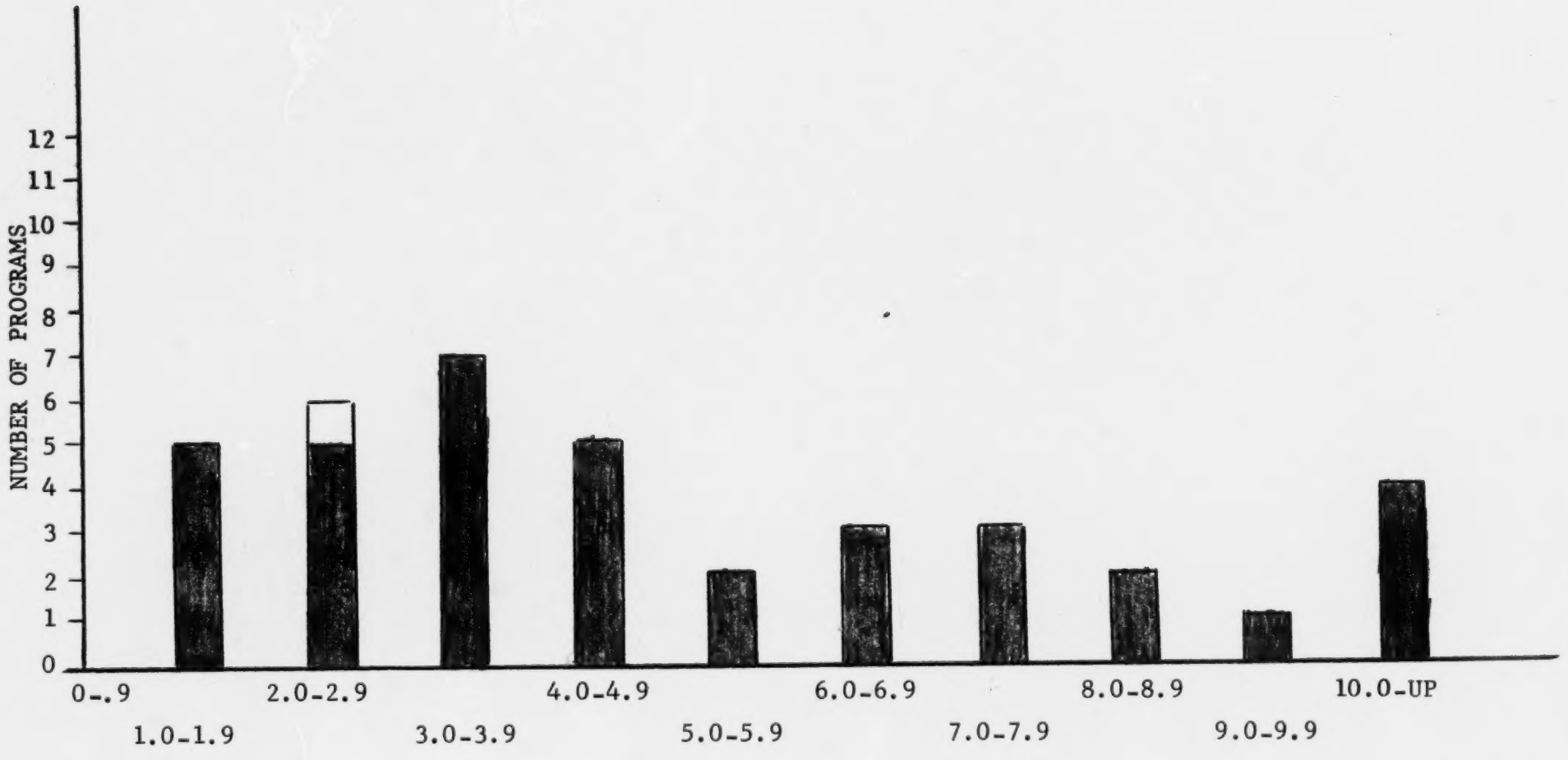

FIG. 2. Frequency distribution of mean error rates for 36 programs previously tested. The non-shaded region locates the value for the test program for this measure at a very low position relative to the other programs tested. (The 36 programs previously tested were not tested by the author under the aegis of the Adult Learning Center.)

indicates that the students learned a considerable amount from the use of the program, consideration may then be given to the more subsidiary criteria, i.e., has the content of the program received the approval of subject-matter experts, are the teacher and student evaluations of the program favorable, and does it have a low error rate? The program that best satisfies these criteria should be the one favored for adoption.

References

McGuigan, F. J. and Peters, R. J., Jr. Assessing the effectiveness of programmed texts -- methodology and some findings. Journal of Programed Instruction, Volume III, No. 1, 1965.

Recommendations for Reporting the Effectiveness of Programed Instruction Materials, Prepared by the Joint Committee on Programmed Instruction and Teaching Machines, Division of Audiovisual Instructional Service, National Education Association, 1966, pp. 1-35, Washington, D. C.

APPENDIX A

STUDENT TIME LOG AND GENERAL INFORMATION FORM

Title of Program _____ Sex _____

Name of Student _____ Race _____

Highest Grade Completed _____ Age _____

Tested Reading Level _____ I. Q. _____

Date & Day	Study Time In Hours And Minutes: Time Started	Time Stopped
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

APPENDIX B

TEACHER EVALUATION FORM

Title of Program _____ Name of Teacher _____

The following questions were designed to help us evaluate the program that you have just gone through with your class. The information that you can furnish will be of great value to us. For each question please check the blank that you feel most adequately describes your opinion. Blank lines have been provided below each question for you to qualify or elaborate your answers. Please feel free to make any comments that will aid us in determining the value of this program.

1. Is the subject-matter of the program academically sound?

_____ Yes

_____ No

_____ Undecided

Comments: _____

_____

_____

2. Was the level of the subject matter appropriate for your class?

_____ Too difficult

_____ Appropriate

_____ Too easy

Comments: _____

_____

_____

3. As contrasted with what you have been able to accomplish with other types of learning material, how much do you feel you were able to get your pupils to learn with this program?

_____ A great deal more than with most other materials.

_____ A little more than with most other materials.

_____ About as much as with other materials.

_____ A little less than with most other materials.

_____ So little as to be a waste of time.

Comments: _____

_____

_____

4. The next time you teach a course in this subject or a similar field, would you:

_____ Prefer to have programs used for at least part of the course?

_____ Prefer not to have programs used?

_____ Not care whether programs are used or not?

Comments: _____

_____

_____

5. To what extent did you enjoy using this program with your class?

_____	_____	_____	_____	_____
Very Unenjoyable	Unenjoyable	50-50	Enjoyable	Very Enjoyable

Comments: _____

_____

_____

6. Do you think this program should be made available for the use of teachers throughout the country?

_____ Yes

_____ No

_____ Don't know

Comments: _____

_____

_____

7. In your own words, would you please summarize your opinion of this program. Include statements about its strong and weak points.

APPENDIX C

STUDENT EVALUATION FORM FOR _____

1. Because a program was used in this course, I believe:

   _____ I learned more.

   _____ it made no difference.

   _____ I learned less.

2. In comparing work done using the program with studying in regular textbooks, I feel that, with the same amount of time and effort:

   _____ I learned much more with the program.

   _____ I learned somewhat more with the program.

   _____ there is no difference.

   _____ I learned somewhat more from studying textbooks.

   _____ I learned much more from studying textbooks.

3. If I were to take another course in this subject or a similar field, I would:

   _____ prefer to have programs used for at least part of the course.

   _____ prefer not to have programs used.

   _____ not care whether programs are used or not.

4. How much do you think you learned from this program?

_____	_____	_____	_____	_____
Learned nothing	Learned a little	Learned a medium amount	Learned quite a bit	Learned very much

5. To what extent did you enjoy going through this program?

_____	_____	_____	_____	_____
Very Unenjoyable	Unenjoyable	50-50	Enjoyable	Very Enjoyable

6. To what extent was the program repetitious?

Much too repetitious	Too repetitious	Moderately repetitious	Slightly repetitious	Not at all repetitious

7. In your own words say what you thought of the program. For example, what did you like about the program? What did you dislike about it, etc.?

ERIC Clearinghouse
MAY 18 1971
on Adult Education